International Journal of Advances in Electronics and Computer Science, ISSN: 2393-2835 Volume-5, Issue-3, Mar.-2018
http://iraj.in

GAZE BASED INTERFACE – A NOVEL APPROACH

KRUPA JARIWALA

Assistant Professor, Department of Computer Engineering, SVNIT, Surat
E-mail: knj@coed.svnit.ac.in

Abstract - The Point and select operations with modern graphical user interfaces (GUIs) greatly enhances the user interaction with the computers. The conventional computer mouse introduces such mapping of hand movement to the on screen cursor with small icons. Gaze based human computer interface is certainly a technology for future computer interaction; therefore, a computationally efficient potential mechanism for the same is of utmost importance. The geometric eye model and edge gradient based eye gaze estimation approach is implemented for the gaze based task selection. It is found possible to estimate the gaze coordinates with good accuracy in normal operating condition and the system can be implemented successfully using ordinary webcams.

Index terms - Eye tracking, gaze estimation, HCI, GUI

I. INTRODUCTION

The modern graphical user interfaces (GUIs) greatly enhances the user interaction with the computers. The conventional computer mouse introduces such mapping of hand movement to the on screen cursor with small icons. Gaze based human computer interface is certainly a technology for the future computer interaction; therefore, a computationally efficient potential mechanism for the same is of utmost importance. There are limited researches which focuses on the challenges webcam-based applications[4] and much of the existing eye tracking researches employ a high resolution camera with a wide angle lenses, which is often set up on the desk below the screen of the computer[1]. Hence, the influence of the eyelid is avoided and high resolution images provide better accuracy. However, general users often set up their webcam on the top of the monitor, therefore an approach proposed by Jariwala[1] that emphasis a general computer user environment for common users with a desktop computer and a low-resolution webcam mounted on the top of the monitor is utilized to establish modern GUI environment.

II. GAZE BASED TASK SELECTION

The Graphical user interface is divided into two operations as pointing and a selection. The cursor is moved to the target area during pointing and when desired location is selected, the activation command is issued. When a traditional input device such as mouse is used as an input, the pointing and selection is explicit. However if gaze is used as the only means of input, then there is a possibility of unintentional activation because the cursor follows the user eyes at all the times. This situation may have negative impact on the interactions.

Eye movements are consequently studied and categorized into fixations and saccades to understand the user's selection objective. Fixations occur when a user is inspecting an object of interest and are characterized by relatively slow eye movements in a limited spatial region[3]. Saccades occur when a user is moving his/her eyes from object to object and are characterized by fast eye movements extending over relatively large spatial regions[3]. While saccades are expected to occur exclusively during pointing, fixations occur both during pointing (e.g. if a user is inspecting an object) and also during selection[3]. Following are the widely used methods to identifying the break between pointing and selection.

A. Dwell Method

This is the most commonly used method to identify inspection or selection fixations by using the time threshold value which can vary from 500 to 1000 ms. This value depend upon system configuration and user preference. The main issue of this method is unintentional activation which is called Midas touch problem[2]. The dwell time is increased to reduce unintentional activation, however, it will also reduce speed and naturalness of interaction. Therefore, setting an optimal dwell time for an individual user is critical, task specific and depends highly on personal preference.

B. Blink and multimodal Methods

Blink and multimodal methods do not use fixation criteria to determine activation. The blink methods check to see that the user closes his/her eyes for longer than a pre-specified time to issue activation. Multimodal methods combine other input methods to issue activation and uses gaze exclusively for pointing.

Here, it is desired to implement gaze alone as an input for an interface design. Therefore, a dwell based method is used to facilitate task selection.

III. EVALUATION MEASURES

The system is setup as proposed by Jariwala[1] using ordinary desktop computer for which camera is

mounted on the top of the computer. Following variations are introduced in the setup (two lighting condition, three distance conditions) so that several other experimental setup are achieved which test different aspects of the system. Screen accuracy is measured to identify the precision of the system throughout the computer screen.

Lighting Condition :

- L1 : Indoor lighting conditions with daylight from the window.
- L2 : Indoor lighting conditions with limited illumination without daylight.

Distance condition :

- D1 : 40 cm between camera and eyes.
- D2 : 50 cm between camera and eyes.
- D3 : 60 cm between camera and eyes.

Screen accuracy :

- Full-screen accuracy (FA) : RMSE over all video frames in the entire screen as shown in figure 1(a).
- Part-screen accuracy (PA) : RMSE over video frames corresponding to gaze in three slices of the screen divided equally (1x3) as PA1, PA2, PA3 as shown in figure 1(b).

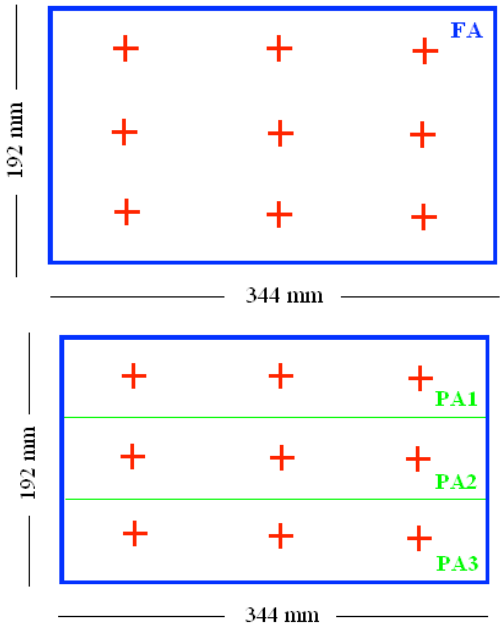

Figure 1 : Screen accuracy (a) Full screen (FA) (b)Part screen (PA1, PA2, PA3).

IV. EXPERIMENTAL RESULTS

System accuracy over varied realistic conditions are characterized. The accuracy is measured in angular degrees as the difference between correct and estimated gaze point. The error between the true and estimated gaze point is measured in root-mean-squared error RMSE.

A 2x3 repeated experiment with varying lighting and distance condition is carried out with participants. Full and part screen accuracy is also calculated to measure the correctness of gaze estimation in the entire screen region.

Based on the feedback from the participants, D1 and D2 position were found most comfortable. No participant found D3 acceptable. Part Screen accuracy (PA1 and PA2) is better and could be used to determine if a user has looked at a moderate size targets. PA3 is worse in all test condition and can not guarantee the knowledge of which part of the screen a user looks at.

Some participants used the system very accurately for which the PA1 accuracy is tracked to under 2° horizontally. This accuracy is closer to commercial gaze tracking systems and is sufficient for applications such as attention analysis.

The decent accuracy in PA1 and PA2 region of the screen is sufficient for a range of applications including the task selections as most operating system menus often appear at the top of the screen and alternatively the system could implement nested task selection once it is noticed. Nevertheless, the accuracy was best during D1 and L1 rather than D2, D3 and L2. The gaze based system can be used for longer session with natural viewing position of D1 in L1 lighting condition for the screen area of PA1 and PA2.

A. Limitations

Though the participants were instructed to use the distance aligned frontal face outlines, they sometimes varied the face orientation, which made the calculation of the estimate difficult. However, the proposed system tolerated minor head pose variations and performed well. In case of major variations in the distance and head pose the system gives wrong estimate and recalibration is required. However, if the face is centered at the screen with reasonably straight alignment, the recalibration can be avoided by saving the training parameters. These parameters can be used later with similar face orientation and distance without any recalibration for the same subject.

V. INTERFACE DESIGN

The interface tool is divided into full-knowledge and no-knowledge tools. Full-knowledge tool rely on the information about the interactive world in which they act to compensate for gaze pointing[3]. They have information about nearby objects, which can help if the target is missed slightly. This method is useful for a custom-designed applications but not useful with mainstream operating systems, as the information about the nearby targets is missing in an interactive environment. Where as, no-knowledge tools do not

need surrounding information for task selection. Therefore, a no knowledge tool is adapted for current implementation of the system.

The PA1 and PA2 region of the screen are used to achieve higher accuracy of the system. The PA3 region is avoided as the accuracy of the system is highly compromised in this region. The layout of the interface tool contains 2x3 grid binned into six targets. The main menu and nested menu is designed to get maximum selection option. The designed targets are expected to interact with the mainstream operating system and invoke appropriate application when activated using dwell based method in no knowledge world where the surrounding information is not available and no anticipatory knowledge is associated with the selection. The targets are activated after appropriate dwell time. The total time taken for the task execution is the time taken combinly for both - pointing and selection.

VI .OBSERVATIONS

On an average the task selection takes approximately 6 to 8 seconds and all participants selected the task correctly in PA1 region. The participants felt in control of the system and described the system as reasonably accurate. All participants reported that they compensated for the system's inaccuracy by overlooking at the extreme corner of the on screen targets. They also adjusted their eye-shape by aligning their face to the camera to better suit the system during the use.

This study was limited for the gaze based task selection without any training. A pre-defined training would make the participants aware about the system and provide more opportunity to compensate for errors. The success rate of the system is reasonably good and acceptable.

CONCLUSIONS

A gaze based interface tool for task selection using moderate size targets is presented, which demonstrated the effectiveness of the approach. The system tracked gaze under 2° horizontally and 2.5^o vertically for D1, L1, PA1. These initial results are promising, and suggest that the proposed approach has potential for cheap consumer devices in near future without any intrusive or specialize equipment.

REFERENCES

[1] K N Jariwala, A Vincent and U D Dalal, "A robust Eye Gaze Estimation using Geometric Eye features", IEEE - third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC-2016), pp. 142–147, 6-8th July 2016.

[2] R. Jacob, "What you look at is what you get: eye movement-based interaction techniques", Proceedings of the Conference on Human Factors in Computing Systems, pp.11-18, 1990.

[3] Henrik H, T. Skovsgaard, J. Mateo, J. P. Hansen, "Evaluating gaze-based interface tools to facilitate point-and-select tasks with small targets", Behavior & Information Technology, Vol. 30, No. 6, pp. 821-831, 2011.

[4] Y. Lin, R. Lin, Yu-Chih Lin, G. Lee, "Real-time eye-gaze estimation using a low-resolution webcam, Multimedia Tools and Applications, Volume 65, Issue 3, pp 543–568, August 2013

[5] Jacob, R.J.K., "Eye tracking in advanced interface design", In the book : Virtual environments and advanced interface design, Oxford University Press, pp. 258–288, 1995.

[6] Sibert, L.E. and Jacob, R.J., "Evaluation of eye gaze interaction", In Proceedings of the Conference on Human Factors in Computing Systems, pp. 281-288, 2000.

[7] Martinez Conde, S. Macknik, and Hubel, D.H., "The role of fixational eye movements in visual perception", Nature reviews. Neuroscience, Vol. 5, No.3, pp. 229–240, 2004.

[8] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision", Cambridge University Press, 2001.

[9] B. Tatler, " Eye movements from Laboratory to life", Current Trends in Eye Tracking Research by Springer, pp. 17-36 , 2013.

[10] B. Knight, M. Horsley and M. Eliot, "Eye Tracking and Learning system", Current Trends in Eye Tracking Research by Springer, pp. 281-286,2013.

★★★